

Feature Extraction and Classification of MRI Using Hybrid RBF Kernel and SVM

Suhas S¹, Dr. C. R. Venugopal²

¹Assistant Professor, Department of Computer Science, The National Institute of Engineering, Mysore, Karnataka, India

²Professor, Department of ECE, Sri Jayachamarajendra College of Engineering, Mysore, Karnataka, India

ABSTRACT

Article Info

Volume 7, Issue 6

Page Number : 418-426

Publication Issue :

November-December-2021

Article History

Accepted : 15 Dec 2021

Published : 30 Dec 2021

An enhanced classification system for classification of MR images using association of kernels with support vector machine is developed and presented in this paper along with the design and development of content-based image retrieval (CBIR) system. Content of image retrieval is the process of finding relevant image from large collection of image database using visual queries. Medical images have led to growth in large image collection. Oriented Rician Noise Reduction Anisotropic Diffusion filter is used for image denoising. A modified hybrid Otsu algorithm termed is used for image segmentation. The texture features are extracted using GLCM method. Genetic algorithm with Joint entropy is adopted for feature selection. The classification is done by support vector machine along with various kernels and the performance is validated. A classification accuracy of 98.83% is obtained using SVM with GRBF kernel. Various features have been extracted and these features are used to classify MR images into five different categories. Performance of the MC-SVM classifier is compared with different kernel functions. From the analysis and performance measures like classification accuracy, it is inferred that the brain and spinal cord MRI classification is best done using MC- SVM with Gaussian RBF kernel function than linear and polynomial kernel functions. The proposed system can provide best classification performance with high accuracy and low error rate.

Keywords: Feature extraction, classification, MRI

I. INTRODUCTION

Support Vector Machine (SVM) is supervised classification technique, which is based on theory of statistical learning. SVM is the best algorithm among algorithms that was developed for pattern

classification. The basic concept of SVM is based on binary classification as it separates data points by a straight line to classify the class label. Whereas, in some datasets, it is not possible to use one straight line to separate the data points. Kernel functions are introduced to overcome the previous issue. What

makes SVM different from other classification algorithms is its outstanding generalization performance. Actually, SVM is one of the few machine learning algorithms to address the generalization problem (i.e., how well a derived model will perform on unseen data). It is not trivial to estimate the generalization error solely based on a training dataset. According to Novikoff's theorem, minimizing the generalization error is equivalent to maximizing the separating margin in support vector classification (SVC).

II. RELATED WORK

Earlier methods used Spatial Gray Level Dependence Method (SGLDM) for feature extraction. Though it was successful, the time consumed by this method was higher and also has higher complexity. To reduce the time for computation A.E.Svalos et al., [6] analyzed the SGLDM method by a pilot application. A technique was proposed by M.Vasantha et al., [7] for extracting the intensity histogram and Gray Level Co-occurrence Matrix (GLCM) features from MR mammogram image. Kernel F-score feature selection (KFFS) method was proposed by Kemal Polat et al., [8] for selecting the features. Results showed that the proposed KFFS functions better than the F-score feature selection. Followed by this, Hsieh-Wei Lee et al., [9] proposed a method for extracting the features from the brain images. This was carried out through integrating the Support Vector Machine (SVM) with the feature selection process in the kernel space. A new method was introduced by Shah et al., [10] in order to elect the salient features for classification and further image processing process. Here, paired t-test was used to eliminate the redundant features and a generic search was employed to detect the hyper-parameter, and to select the salient features.

A texture based classification system was proposed by Sidhuet al., [13] using SVM and wavelet transform. The core concept of this classification system was to identify and analyze the factors that considerably affect the performance of SVM and wavelet transform during the process of texture classification. A batch type learning vector quantization technique for segmentation was proposed by Miin-Shen Yang et al., [12]. It provides good accuracy and quality for the accurate measurement of hippocampus volume in MR images. The methodology was compared with the generalized Kohonen's competitive learning method.

III. PROPOSED WORK

Classification is the most widely used strategy used for recognizing the objects based on its features. Figure 2 shows the classification model proposed by us to identify the different types of images present in the image set. Two major parts of the model are feature extraction and classification. The textural features obtained from the MR images are given as input to the classification part of the model. The classifier used is the Support Vector Machine (SVM) with Gaussian RBF kernel.

A. Feature Extraction

The margins must be set Various techniques for extracting features from MRI brain images have been reported in the literature, the most common are: Discrete Wavelet Transform (DWT) [9], Gabor filters [10] and Gray Level Co-occurrence Matrix (GLCM) [11]. Both DWT and Gabor Filter methods produce feature vectors with a large number of elements which necessitates the use of size reduction techniques prior to feeding the feature vectors to the classifier. On the other hand, The Gray Level Co-occurrence Matrix (GLCM) has proven to be superior in terms the dimension of the feature vectors and thus is more appropriate for MRI image classification.

GLCM is a statistical technique for extracting texture features from images [11]. It assumes that the texture of normal tissues is very different from the texture of tumor tissues. The texture features extracted from the GLCM matrix are: contrast, correlation, energy, homogeneity. Selecting a good set of features improve the process of classification. Additional second order features were also extracted from this matrix which are: mean, standard deviation, entropy, root mean square, variance, kurtosis, skewness, AutoCorrelation, Maximum probability ,Sum average, Sum of squares: Variance, Inverse difference normalized, Sum entropy, Sum Variance, Difference variance, Information measure of correlation . All features used in this work are listed below. These features were extracted from different brain and spinal cord MR Images. The averaged features values are shown in Table 1.

Contrast (*Con*). Contrast is a measure of intensity of a pixel and its neighbor over the image, and it is defined as

$$Con = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x - y)^2 f(x, y)$$

Correlation

$$Corr = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x, y) f(x, y) - M_x M_y}{\sigma_x \sigma_y}$$

Autocorrelation is a measure of the magnitude of the fineness and coarseness of texture.

$$Autocorrelation = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) ij$$

Energy (*En*). Energy can be defined as the quantifiable amount of the extent of pixel pair repetitions. Energy is a parameter to measure the similarity of an image. Energy is also referred to as angular second moment, and it is defined as

$$En = \sqrt{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f^2(x, y)}$$

Homogeneity or Inverse Difference Moment is a measure of the local homogeneity of an image. IDM may have a single or a range of values so as to determine whether the image is textured or nontextured

$$IDM = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \frac{1}{1 + (x - y)^2} f(x, y)$$

Mean

$$M = \frac{1}{(m \times n)} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y)$$

Standard deviation

$$SD(\sigma) = \sqrt{\frac{1}{(m \times n)} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (f(x, y) - M)^2}$$

Entropy is calculated to characterize the randomness of the textural image and is defined as

$$E = - \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y) \log_2 f(x, y)$$

Coarseness (*Cness*). Coarseness is a measure of roughness in the textural analysis of an image. For a fixed window size, a texture with a smaller number of texture elements is said to be coarser than the one with a larger number. The rougher texture means higher coarseness value. Fine textures have smaller values of coarseness. It is defined as

$$Cness = \frac{1}{2^{m+n}} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y)$$

Variance

$$Var = \left(\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} |x - y| f(x, y) \right)$$

Homogeneity

$$homogeneity = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + |i - j|}$$

Kurtosis . The shape of a random variable's probability distribution is described by the parameter

called Kurtosis. For the random variable X , the Kurtosis is denoted as $Kurt(X)$ and it is defined as

$$Kurt(X) = \left(\frac{1}{m \times n}\right) \frac{\sum(f(x, y) - M)^4}{SD^4}$$

Skewness (Sk). Skewness is a measure of symmetry or the lack of symmetry. The skewness of a random variable X is denoted as (X) and it is defined as

$$Sk(X) = \left(\frac{1}{m \times n}\right) \frac{\sum(f(x, y) - M)^3}{SD^3}$$

Cluster Prominence is a measure of the skewness and asymmetry of the GLCM. A higher values implies more asymmetry about the mean while a lower value indicates a peak near the mean value and less variation about the mean.

cluster prominence

$$= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^4 p(i, j)$$

Cluster Shade is a measure of the skewness and uniformity of the GLCM. A higher cluster shade implies greater asymmetry about the mean.

$$\text{cluster shade} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^3 p(i, j)$$

Maximum Probability is occurrences of the most predominant pair of neighboring intensity values.

$$\text{maximum probability} = \max(p(i, j))$$

Sum of Squares or Variance is a measure in the distribution of neighboring intensity level pairs about the mean intensity level in the GLCM.

$$\text{sum squares} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 p(i, j)$$

There are four aspects of feature extraction: • feature construction; • feature subset generation (or search strategy); • evaluation criterion definition (e.g. relevance index or predictive power); • evaluation criterion estimation (or assessment method). The last three aspects are relevant to feature selection. Filters and wrappers differ mostly by the evaluation criterion. It is usually understood that filters use criteria not involving any learning machine, e.g. a relevance index based on correlation coefficients or test statistics, whereas wrappers use the performance of a learning machine trained using a given feature subset.

B. Classification

Classification is the next step after feature extraction and it is a supervised learning procedure. It involves two steps training and testing. During the training phase, the classifier is trained with features from training images. In testing phase, an unknown image's features are given to the classifier and it has to classify the image as 'tumor affected' or 'tumor not affected'.

SVMs are the most well-known learning systems based on kernel methods. First introduced by Vapnik [16], it is as an alternative to neural networks, and that has been successfully employed to solve clustering problems, especially in biological applications. It performs classification by constructing an N-dimensional hyperplane that optimally separates the data into two categories. A classification task usually involves training and testing data which consist of some data instances. Each instance in the training set contains one "target value" (class labels) and several "attributes" (features). The goal of SVM is to produce a model, which predicts target value of data instances in the testing set, which are given only

the attributes. Given a training set of instance-label pairs

Here training vectors x_i are mapped into a higher (maybe infinite) dimensional space by the function Φ . Then SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space. $C > 0$ is the penalty parameter of the error term. Furthermore, $K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j)$ is called the kernel function. The kernel is used to transform data from the input (independent) to the feature space. Following are the four basic types of kernel functions:

Linear Kernel: The Linear kernel is the simplest kernel function. It is given by the inner product $\langle x, y \rangle$ plus an optional constant c . Kernel algorithms using a linear kernel are often equivalent to their non-kernel counterparts. The benefit of the linear kernel is that it is incredibly simple and only has the constant term c as a parameter. The linear kernel is typically used on data sets with large amounts of features as increasing the dimensionality on these data set does not necessarily improve separability. It is defined as

$$K(x, y) = x^T y + c$$

Polynomial Kernel: The Polynomial kernel is a non-stationary kernel, does involve taking the inner product from a higher dimension space. Adjustable parameters are the slope α , the constant term c and the polynomial degree d . It is defined as

$$K(x, y) = (\alpha x^T y + c)^d$$

Gaussian Radial Basis Kernel function (GRBF): It is defined as

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right)$$

Exponential Kernel: The exponential kernel is closely related to the Gaussian kernel, with only the square of the norm left out. It is also a radial basis function kernel. It is defined as

$$K(x, y) = \exp\left(-\frac{\|x - y\|}{\sigma}\right)$$

Anova Kernel: The ANOVA kernel is also a radial basis function kernel, just as the Gaussian and Laplacian kernels. It is defined as

$$K(x, y) = \sum_{K=1}^n \exp(-(x^K - y^K)^2)$$

The choice of a Kernel depends on the problem at hand because it depends on what we are trying to model. A polynomial kernel, for example, allows us to model feature conjunctions up to the order of the polynomial. Radial basis functions allows picking out circles (or hyperspheres) – in contrast with the linear kernel, which allows only picking out lines (or hyperplanes).

Support vector machines (SVMs) of both linear and Gaussian radial basis function (GRBF) (non-linear kernel) are found to be computationally inexpensive

Classification is a data mining (machine learning) technique used to predict group membership for data instances Classification is a problem of detecting the classes of data with the help of some already known classes. This is also called as supervised classification Thus the requirement is that new individual items are placed into groups based on quantitative information on one or more measurements, traits or characteristics, etc. and based on the training set in which previously decided groupings are already established. On the other hand, the classification in which no expert is present for prediction is called as Unsupervised Classification.

The pseudocode for classification of the MRI using SVM to evaluate the kernel function is as shown below.

Algorithm: Classification using SVM with GRBF

Input: training data set T, constant $c > 0$ for tuning errors and margin size

Training:

1. Start
2. Input the dataset
3. Read the input images
3. Get all unique image categories and the number N of categories in the Training set
4. Apply the SVM Machine learning with four kernel functions (linear, polynomial, Exponential and Gaussian Radial Based Function (RBF) & ANOVA) and extract the features from the input image.
5. Classify the input image into the appropriate category based on the input data
6. If obtained Accuracy and validity is NOT acceptable then go to step 4
7. End

The proposed algorithm takes the dataset as input, classifies it, and applies the SVM with four kernels to

specify the Hyperplane and the category. Then it checks the obtained accuracy and validity.

IV. EXPERIMENTAL RESULTS

The test set for this evaluation experiment were run on MR Images procured from hospital. Matlab 9.3 software platform is use to perform the experiment. The PC for experiment is equipped with an AMD FX 6300 Quad Core 3.5GHz processor and 16GB memory. Around 130 MR Images of different sequences ranging from different views of the brain and spinal cord were considered for the experiment ranging from T1 weighted, T2 weighted, STIR, DWI, ADC and TOF. The images were all pre-processed, denoised, segmented, and used here for feature extraction and classification. Different kernels with Multi Class SVM were used to classify the images into 5 different types based on the type of the image. The proposed algorithm takes the dataset as input, applies the SVM with four kernels to specify the hyperplane, and then classifies it. Then it is checked for accuracy.

TABLE I
FEATURES EXTRACTED FROM THE DIFFERENT SETS OF MR IMAGES AND THEIR VALUES

Feature	Brain Set 1	Brain Set 2	Brain Set 3	Spinal Cord set 4	Spinal Cord set 5
Contrast.	0.7961	0.7320	0.4065	0.3640	0.44139
Correlation .	0.9245	0.9566	0.9522	0.9627	0.9366
Cluster Prominence.	901.906	2186.85	889.89	1535.27	681.99
Cluster Shade.	82.638	168.28	84.345	135.978	68.205
Dissimilarity.	0.2343	0.2286	0.1540	0.1300	0.1604
Energy.	0.4423	0.4389	0.4840	0.5634	0.5037
Entropy.	1.2780	0.9399	1.1562	1.0261	1.0875
Homogeneity /Inverse difference moment.	0.9284	0.9284	0.9555	0.9570	0.9559
Maximum probability.	0.6356	0.6304	0.6732	0.7404	0.6848
Sum of squares: Variance.	11.9474	17.260	9.499	9.5830	7.991
AutoCorrelation.	11.6303	17.5325	9.368	9.462	7.8184

The performance of the proposed method has been evaluated in terms of accuracy. Among the 130 MR images, there were 5 types of images with 3 different views of the brain and 2 different views of the spinal cord as shown in figures 1 to 5 below. The Brain MRI sets were three sets of 30 each while the spinal cord sets were two sets of 20 each These were used for training purposes and classified using MC-SVM and different kernel functions. The results show that MC-SVM with Gaussian Radial Basis function kernel gives the highest classification accuracy from among all the tested kernels.

$$\text{accuracy} = \frac{\text{items classified correctly}}{\text{all items classified}}$$

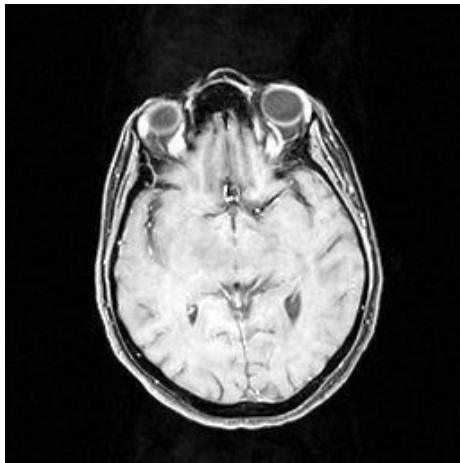


Figure 1. T1 weighted MRI sequence (Brain set 1)

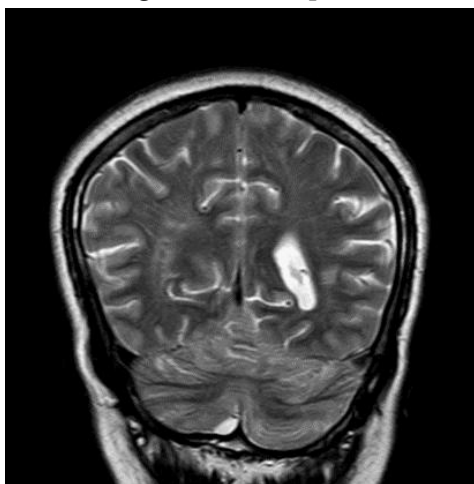


Figure 2. T2 weighted MRI sequence (Brain set 2)



Figure 3. Time of flight angiography (Brain set 3)



Figure 4. Inversion recovery MRI sequence (Spinal cord set 4)



Figure 5. Diffusion weighted (Spinal cord set 5)

TABLE II. CLASSIFICATION OF THE DIFFERENT MR IMAGES BASED ON THE KERNELS AND ITS ACCURACY

Type of kernel	No. of images used for testing	Images classified					Percentage of correct classification (accuracy)
		Brain MRI set 1	Brain MRI set 2	Brain MRI set 3	Spinal cord MRI set 1	Spinal cord MRI set 2	
Linear	130	25	28	27	18	17	88.46
Polynomial		28	27	27	18	17	90.00
GRBF		29	30	30	20	20	99.23
Exponential		29	29	29	19	18	95.38
ANOVA		28	29	29	19	19	95.38

V. CONCLUSION AND DISCUSSIONS

In this study, our aim was to develop a Multi-Class SVM with kernel method to classify different sequences of MRI namely T1-weighted, T2-weighted, Diffusion weighted, Time of flight images. Different features were extracted from the images and the classification accuracy of the MC-SVM with GRBF kernel outperformed all other kernels with a classification accuracy of 99.23%

VI. REFERENCES

[1]. Hatice Cinar Akakin ; Metin N. Gurcan (2012). Content-Based Microscopic Image Retrieval System for Multi-Image Queries, IEEE Transactions On Information Technology In Biomedicine, Vol. 16, No. 4, July 2012

[2]. Xiaofan Zhang; Wei Liu; Sunil Badve; Towards Large-Scale Histopathological Image Analysis: Hashing-Based Image Retrieval (2013). IEEE Transactions on Medical Imaging.

[3]. Madina Hamiane ; Fatema Saeed; SVM Classification of MRI Brain Images for Computer Assisted Diagnosis(2017). International Journal of Electrical and Computer Engineering (IJECE) Vol. 7, No. 5, October 2017, pp. 2555~2564

[4]. Jayashri Joshi; A. C. Phadke; Feature Extraction and Texture Classification in MRI(2012).

International Journal of Computer and Communication Technology, Vol 3, No.2

[5]. Icke, Guzide, Kamarthi, V.Sagar , “Feature extraction through discrete wavelet transform coefficients”, Intelligent Systems in Design and Manufacturing VI, Proceedings of the SPIE, Volume 5999, pp. 27-35, 2005. [10]

[6]. Svolos A E, Todd-Pokropek A. Time and space results of dynamic texture feature extraction in MR and CT image analysis. IEEE transactions on journal of information technology in Biomedicine 1998; 2(2): 48-54.

[7]. Vasantha M, Subbiah Bharathi V, Dhamodharan R. Medical Image Feature, Extraction, Selection and Classification. International Journal of Engineering Science and Technology 2010; 2(6): 2071-2076.

[8]. Kemal Polat, Salih Gunes. A new feature selection method on classification of medical datasets: Kernel F-score feature selection. International Journal of Expert Systems with Applications 2009; 36(7): 10367-10373.

[9]. Hsieh-Wei Lee, King-Chu Hung, Bin-Da Liu, Sheau-Fang Lei, Hsin-Wen Ting. Realization of High Octave Decomposition for Breast Cancer Feature Extraction on Ultrasound Images. IEEE transactions on Circuits and System 2011; 58(6): 1287-1299.

[10]. S. Shah, & N Chauhan, “Classification of Brain MRI Images using Computational Intelligent

- Techniques”, International Journal of Computer Applications, Volume 124 , No.14. 2015.
- [11]. Zulpe, N. & Pawar, V., “GLCM Textural Features for Brain Tumor Classification”, International Journal of Computer Science Issues, Vol. 9, Issue 3, No 3, 2012 [12]
- [12]. V. Josephine, & P. Latha, “SVM -Based Automatic Medical Decision Support System For Medical Image”, Journal of Theoretical and Applied Information Technolog, Vol. 66 No.3. 2014. [13]
- [13]. A Yuniarti, "Classification and numbering of dental radiographs for an automated human identification system,” TELKOMNIKA Telecommunication, Computing, Electronics and Control. vol. 10, no. 1, pp. 137-146, 2012.
- [14]. A. Kulkarni, & R. Kamathe, “MRI Brain Image Segmentation by Edge Detection and Region Selection”, International Journal of Technology and Science, Issue. 2, Vol. 1, 2014
- [15]. Yang, Miin-Shen & Lin, Karen & Liu, Hsiu-Chih & Lirng, Jiing-Feng. (2007). Magnetic resonance imaging segmentation techniques using batch-type learning vector quantization algorithms. Magnetic resonance imaging. 25. 265-77.
- [16]. S Sidhu, K Raahemifar. Texture classification using wavelet transform and support vector machines. Electrical and Computer Engineering Canadian Conference proceeding 2005; 941-944

Cite this article as :

Suhas S, Dr. C. R. Venugopal, "Feature Extraction and Classification of MRI Using Hybrid RBF Kernel and SVM", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 6, pp. 418-426, November-December 2021. Available at
doi : <https://doi.org/10.32628/CSEIT2176104>
Journal URL : <https://ijsrcseit.com/CSEIT2176104>